

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

КОТВИЦЬКИЙ ІГОР ВАЛЕРІЙОВИЧ



УДК 621.391.83

**ОЦІНЮВАННЯ ЯКОСТІ СПОТВОРЕНИХ
МОВНИХ ТА МУЗИЧНИХ СИГНАЛІВ**

Спеціальність 05.09.08 – прикладна акустика та звукотехніка

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
кандидата технічних наук

Київ – 2020

Дисертацією є рукопис

Робота виконана на кафедрі акустики та акустoeлектроніки Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» Міністерства освіти і науки України.

Науковий керівник – доктор технічних наук, професор

Продеус Аркадій Миколайович

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря

Сікорського» (м. Київ), професор кафедри акустики та акустoeлектроніки.

Офіційні опоненти – доктор технічних наук, професор

Мислович Михайло Володимирович

Інститут електродинаміки Національної академії наук

України (м. Київ), відділу теоретичної електротехніки та діагностики електротехнічного обладнання

– кандидат технічних наук, доцент

Мелешко Микола Андрійович

Національний авіаційний університет, професор кафедри комп'ютерних мультимедійних технологій

Захист відбудеться «22» вересня 2020 р. о 14:30 годині на засіданні спеціалізованої вченої ради Д26.002.19 у Національному технічному університеті України «Київський політехнічний інститут ім. Ігоря Сікорського» за адресою 03056, м. Київ-56, просп. Перемоги, 37, корпус 12, аудиторія 412.

З дисертацією можна ознайомитись у бібліотеці Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» за адресою 03056, м. Київ-56, просп. Перемоги, 37.

Автореферат розісланий «21» серпня 2020 р.

Вчений секретар
спеціалізованої вченої ради



В.Б. Швайченко

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Оцінювання якості спотворених мовних та музичних сигналів є основною складовою акустичної експертизи комунікаційних каналів. Перевагою суб'єктивного оцінювання якості сигналів є еталонний характер отриманих результатів, оскільки якість звукових сигналів оцінюють кінцеві користувачі. Суттєвим недоліком суб'єктивного оцінювання є високі вимоги до людських, часових та фінансових ресурсів. Такого недоліку позбавлені об'єктивні (інструментальні) методи й системи оцінювання, чим і пояснюється незмінний інтерес науковців та інженерів до таких систем. Значний внесок в розвиток об'єктивних методів оцінювання якості акустичних сигналів зробили такі вчені як Покровський Н.Б., Биков Ю.С., Сапожков М.А., Калінцев Ю.К., Фант Г., Рабінер Л., Янг В., Цвіккер Е., Фельдкеллер Р., Луїзоу П., Мюллер С., Коте Т., Квакенбуш І., Бірендс Й., Блауерт Й., Бредлі Д., Арвейлер І. та інші.

Суттєвим ускладненням при застосуванні методів об'єктивного оцінювання є різноманітність причин спотворення акустичних сигналів. Це дія шумових та ревербераційних завад, нерівномірність амплітудної частотної характеристики та обмеженість смуги частот тракту, нелінійність фазочастотної характеристики тракту, кліпування сигналів, спотворення сигналів при використанні складних алгоритмів їх обробки тощо.

Іншим ускладненням є надзвичайно велика кількість об'єктивних мір якості сигналів, що значно ускладнює вибір потрібної міри. Однією з причин такого різноманіття є бажання якнайкраще узгодити результати суб'єктивного та об'єктивного оцінювання, що призводить до ускладнення об'єктивних мір якості. При цьому, на жаль, залишаються не до кінця дослідженими потенційні можливості простих для обчислень об'єктивних мір, таких, зокрема, як сегментне відношення сигнал-шум та лог-спектральні спотворення.

Нагальним завданням є отримання еталонних результатів суб'єктивного оцінювання якості та розбірливості мови – від наявності та достовірності таких еталонів залежить принципова можливість практичного використання об'єктивних мір якості та розбірливості мови. Розв'язання цього завдання потребує обробки великих обсягів звукового матеріалу та автоматизації процедур суб'єктивних випробувань. Тому актуальним є питання створення простих, недорогих та надійних систем автоматизації суб'єктивного оцінювання якості та розбірливості мови.

Оскільки оцінювання розбірливості мови об'єктивними методами є набагато складнішим, порівняно з оцінюванням якості мовних сигналів, актуальним є питання про пошук умов, за яких можливе використання простих мір якості замість складних мір розбірливості мови.

Недостатньо вивченим на теперішній час є також питання суб'єктивного та об'єктивного оцінювання якості музичних сигналів. Праці таких дослідників як Алдошина І., Прітте Р., Гянакопулос Т., Пікракіс А. переважно стосуються оцінювання параметрів музичних сигналів, а питання об'єктивного оцінювання якості спотворених музичних сигналів в них майже не висвітлено. Цей недолік

в певній мірі компенсовано в працях Бірендс Д., Тіеде Т., Хюбер Р., Барбеда Д. та інших, де запропоновано так звані «перцепційні міри» якості акустичних сигналів, що містять різні моделі слухової системи людини. Складність таких мір та високі вимоги до обчислювальних ресурсів є їхнім суттєвим недоліком, що мотивує до пошуків більш простих, і водночас надійних, технічних рішень.

Оскільки оцінювання якості мовних та музичних сигналів є основою сучасних цифрових систем акустичної експертизи комунікаційних каналів, створення відповідного математичного та програмного забезпечення є **актуальною** науково-технічною задачею, що має важливе прикладне значення. Цим визначається актуальність та практичне значення теми дисертаційного дослідження.

Зв'язок роботи з науковими планами, програмами, темами.

Дисертаційну роботу виконано на кафедрі акустики та акустoeлектроніки Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» в рамках НДР «Об'єктивне та суб'єктивне оцінювання якості та розбірливості мовлення» (№ДР 0116U008959).

Мета та задачі дослідження. Метою роботи є вдосконалення методів суб'єктивного та об'єктивного оцінювання якості мовних і музичних сигналів, підданих лінійним та нелінійним спотворенням.

Для досягнення цієї мети необхідно знайти розв'язок наступних завдань:

1. Зіставлення потенційних можливостей об'єктивних мір якості мовних та музичних сигналів, що дозволяє здійснювати обґрунтований вибір кращої міри за такими показниками як достовірність та швидкість акустичної експертизи.
2. Поглиблений аналіз маловідомих особливостей та виявлення невідомих властивостей простих для обчислень мір якості акустичних сигналів, що забезпечують високу швидкодію обчислень.
3. Визначення умов, за яких можливе використання мір якості мовних сигналів, спотворених шумовою завадою, замість мір розбірливості мови, що дозволяє спростити процедуру оцінювання без втрати її достовірності.
4. Формування низки еталонних характеристик, що характеризують чутливість слухової системи людини до лінійних та нелінійних спотворень сигналів в трактах передачі й перетворення акустичної інформації, шляхом автоматизованого суб'єктивного оцінювання якості мовних і музичних сигналів, а також розбірливості мовних сигналів.
5. Зіставлення потенційних можливостей алгоритмів просторово-часової обробки акустичних сигналів в мікрофонних масивах в задачі притлумлення шумової та ревербераційної завад за умови збереження високої якості акустичних сигналів.

Об'єктом дослідження є процеси обробки мовленнєвих та музичних сигналів, спрямовані на визначення якості сигналів, спотворених в каналах передачі акустичної інформації.

Предметом дослідження є методи обробки спотворених сигналів в системах суб'єктивної та об'єктивної акустичної експертизи.

Методи дослідження. При виконанні досліджень використано низку методів постановки і розв'язання задач обробки акустичних сигналів. Так, для опису процесів оцінювання якості мовних та музичних сигналів використано аналітичні та комп'ютерні моделі слухової системи людини, методи теорії цифрової обробки сигналів та методи оцінювання якості акустичних сигналів. Для визначення умов, за яких можливо використання мір якості акустичних сигналів замість мір розбірливості мови, використано методи об'єктивного та суб'єктивного оцінювання розбірливості мови. При зіставленні алгоритмів обробки акустичних сигналів в мікрофонних масивах використано методи просторово-часової обробки акустичних сигналів. Достовірність отриманих результатів підтверджується зіставленням одержаних результатів з опублікованими в науковій літературі відповідними результатами, в тому числі експериментальними.

Наукова новизна одержаних результатів.

1. Вперше для реальних мовних та музичних сигналів встановлено, що слухова система людини є чутливою до знаку різниці часу групової затримки для низьких та високих частот, а також отримано суб'єктивні та об'єктивні оцінки ступеню впливу нелінійності фазової характеристики тракту на якість сигналів.

2. Вперше запропоновано використання ексцесу та його функціональних перетворень як об'єктивних мір якості мовних та музичних сигналів, спотворених кліпуванням.

3. Вперше вказано на підвищену чутливість до спектральних властивостей сигналів простої в обчислювальному плані об'єктивної міри якості сигналів у вигляді лог-спектральних спотворень.

4. Знайшло подальший розвиток положення щодо підвищеної чутливості сегментного відношення сигнал-шум до похибки вирівнювання сигналів та запропоновано рекомендації зі зниження такої чутливості.

5. Знайшло подальший розвиток зіставлення, шляхом об'єктивного та суб'єктивного оцінювання якості звуку, алгоритмів обробки музичних сигналів в багатомікрофонних масивах.

Практичне значення одержаних результатів полягає в наступному:

1. Встановлено принципову можливість використання, при акустичній експертизі комунікаційних каналів, порівняно простих, з обчислювальної точки зору, об'єктивних мір якості сигналів SSNR, LSD та BSD замість складних перцепційних мір PESQ та PEAQ.

2. Встановлено принципову можливість використання коефіцієнту ексцесу та його простих функціональних перетворень як об'єктивних мір якості мовних та музичних сигналів, спотворених кліпуванням, що дозволяє спростити та підвищити надійність процедури виявлення кліпування.

3. Визначено межі можливого застосування порівняно простих, з обчислювальної точки зору, методів оцінювання якості мовних сигналів замість більш складних методів оцінювання розбірливості мови.

4. Розроблено комплекс комп'ютерних програм для автоматизації артикуляційних випробувань каналів передачі української мови в умовах дії шумової завади та реверберації, що дозволяє значно знизити вимоги до фінансових, часових та людських ресурсів при проведенні артикуляційних випробувань.

5. Виконано порівняння двох алгоритмів просторово-часової обробки, таких як DS та LCMV, за критерієм якості музичного сигналу, спотвореного адитивними завадами. Показано, що простіший для обчислення алгоритм DS дозволяє отримати якісніші сигнали.

Результати дисертаційної роботи впроваджені у практику оцінювання якості звукоізоляції шляхом оцінювання розбірливості мови у вигляді математичного забезпечення систем оцінювання розбірливості мови для науково-виробничого підприємства «ТОВ Глобал Акустікс», а також в учбовий процес у курсах лекцій кафедри акустики та акустoeлектроніки з дисципліни «Комп'ютерна обробка акустичних сигналів» у розділах «Оцінювання якості акустичних сигналів», «Автоматизація оцінювання розбірливості мови» та дисципліни «Захист акустичної інформації» у розділі «Прогнозування та вимірювання розбірливості мови» (№ДР 0116U008959).

Впровадження результатів дисертаційної роботи підтверджено відповідними актами.

Особистий внесок здобувача відображений публікаціями [1-16]. У наукових працях, виконаних у співавторстві, здобувачу належить: в роботах [1, 10] комп'ютерне моделювання фазових спотворень, проведення експерименту та аналіз отриманих результатів, в роботах [2, 4, 6, 7, 9, 11, 13, 14] – організація експерименту та аналіз результатів експерименту, в роботах [3, 5, 8, 12, 15, 16] – дослідження комп'ютерної моделі системи артикуляційних досліджень та проведенні експерименту.

Апробація результатів дисертації. Основні положення та результати дисертації обговорювались на:

- V Міжнародній науково-технічній конференції молодих вчених «Інформаційні проблеми теорії акустичних, радіoeлектронних і телекомунікаційних систем» (IPST-2016), 29.11-1.12, 2016, Харків;

- Міжнародній науково-технічній конференції молодих вчених «International Young Scientists Forum on Applied Physics and Engineering (YSF)», Львів, 17.10.2017;

- V International Conference "Methods and Systems of Navigation and Motion Control" (MSNMC-2018), October 16-18, 2018, Kyiv, Ukraine;

- IV-Vth International Conferences Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2017, 2019, Kyiv, Ukraine;

- 37-39th International Conferences on Electronics and Nanotechnology (ELNANO), 2017, 2018, 2019, Kyiv, Ukraine;

- 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), February 20-24, 2018, Lviv-Slavske, Ukraine;

- 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), October 2019, Kyiv, Ukraine.

Публікації. За результатами досліджень опубліковано 16 наукових праць, у тому числі 5 статей у наукових фахових виданнях України, які включені до міжнародних наукометричних баз (Index Copernicus, Google Scholar), 1 свідоцтво про авторське право на комп'ютерну програму, 10 тез доповідей в збірках матеріалів конференцій.

Структура та об'єм дисертаційної роботи. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та двох додатків. Загальний обсяг роботи становить 161 сторінок, у тому числі 116 сторінок основного тексту, 79 рисунків та списку використаних джерел із 77 найменувань.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано вибір теми дослідження, актуальність дисертаційної роботи, сформульовано мету і задачі дослідження, визначено наукову новизну і практичну цінність отриманих результатів, наведено дані щодо їх апробації.

У **першому розділі** проведено аналіз сучасного стану акустичної експертизи каналів передачі, запису та відтворення мовних та музичних сигналів, спотворених шумом, реверберацією та дією низки факторів, таких як обмеження смуги частот, фазові спотворення, кліпування (рис. 1).

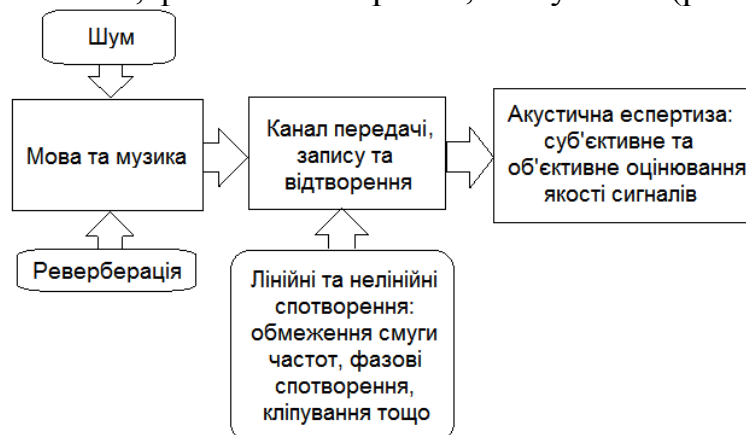


Рис. 1. Задача акустичної експертизи мовних та музичних сигналів

Перевагою суб'єктивного оцінювання якості сигналів є еталонний характер отриманих результатів, а суттєвим недоліком є високі вимоги до людських, часових та фінансових ресурсів. Тому пошук простих для обчислення та достовірних об'єктивних методів акустичної експертизи є актуальним завданням.

Оскільки отримання достовірних результатів оцінювання потребує обробки великих обсягів звукового матеріалу, це робить актуальним питання створення простих та надійних автоматизованих систем оцінювання якості та розбірливості мови.

Значним ускладненням при застосуванні методів об'єктивного оцінювання якості акустичних сигналів є надзвичайно велика кількість об'єктивних мір якості. Тому актуальним є питання поглибленого вивчення властивостей простих для обчислень об'єктивних мір.

Оскільки оцінювання розбірливості мови об'єктивними методами є набагато складнішим, порівняно з оцінюванням якості мови, виникає питання про пошук умов, за яких можливе використання простих мір якості замість складних мір розбірливості мови.

Одним із найбільш сучасних підходів до реєстрації звукового поля є використання мікрофонних масивів, що допомагають зменшити вплив шумової та ревербераційної завад. Проте на сьогодні невизначеним є питання про доцільність використання досить складного алгоритму LCMV (алгоритм мінімізації відхилення із лінійними обмеженнями) замість простого алгоритму DS (алгоритм затримки та підсумовування) при записах мови та музики.

Вказані вище актуальні питання зумовили постановку задач, що розв'язуються у даній дисертації.

У **другому розділі** представлено результати порівняння об'єктивних та суб'єктивних оцінок якості мовних та музичних сигналів при лінійних та нелінійних спотвореннях. При цьому розглянуто спотворення у вигляді обмеження смуги частот, нелінійності фазочастотної характеристики каналу та кліпування (обмеження амплітуди).

Узагальнену схему досліджень представлено на рис. 2, згідно якій як при суб'єктивному, так і при об'єктивному методах оцінювання якості сигналів виконується порівняння спотвореного сигналу із еталонним чистим сигналом.

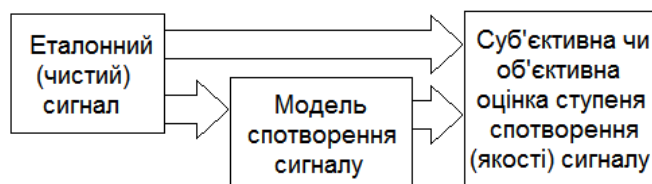


Рис. 2. Узагальнена схема оцінки якості спотворених мовних та музичних сигналів

При суб'єктивному оцінюванні якості сигналів автором використано п'ятибальну шкалу Degradation Mean Opinion Score (DMOS), згідно якої оцінюється ступінь деградації якості спотвореного сигналу відносно еталонного сигналу.

При об'єктивному оцінюванні якості сигналів автором використовувалися такі міри якості як сегментне відношення сигнал-шум (SSNR):

$$SSNR = \frac{1}{M} \sum_{m=1}^M 10 \lg \left[\frac{\sum_{n=R(m-1)+1}^{Rm} x^2(n, m)}{\sum_{n=R(m-1)+1}^{Rm} [x(n, m) - y(n, m)]^2} \right], \quad (1)$$

лог-спектральні спотворення (LSD):

$$LSD = \frac{2}{KL} \sum_l \sum_{k=0}^{\frac{K}{2}-1} |G\{X(l, k)\} - G\{Y(l, k)\}|, \quad (2)$$

$$\text{де } G\{X(l, k)\} = \max \{20 \lg(|X(l, k)|), \delta\}, \quad \delta = \max_{l, k} \{20 \lg(|X(l, k)|)\} - 50,$$

барк-спектральні спотворення (BSD):

$$BSD = \frac{\sum_l \sum_{k=0}^{\frac{K}{2}-1} [B\{X(l, k)\} - B\{Y(l, k)\}]^2}{\sum_l \sum_{k=0}^{\frac{K}{2}-1} [B\{X(l, k)\}]^2}, \quad (3)$$

частотно-зважене сегментне відношення сигнал-шум (fwSSNR):

$$fwSSNR = \frac{10}{M} \sum_{m=1}^M \frac{\sum_{k=1}^K W(k, m) \lg \frac{|X(k, m)|^2}{(|X(k, m)| - |Y(k, m)|)^2}}{\sum_{k=1}^K W(k, m)}, \quad (4)$$

де $x(n, m)$ і $y(n, m)$ – m -та вибірка n -го фрейму вхідного і вихідного сигналів фільтра $x(n)$ та $y(n)$, відповідно; $X(l, k)$ та $Y(l, k)$ – амплітудні спектри l -го фрейму сигналів $x(n)$ та $y(n)$, відповідно; $B\{X(l, k)\}$ та $B\{Y(l, k)\}$ – барк-спектри l -го фрейму сигналів $x(n)$ та $y(n)$, відповідно; k – номер критичної смуги; K – кількість критичних смуг; $|X(k, m)|$ – амплітудний спектр чистого сигналу; $W(k, m)$ – вагові коефіцієнти.

Перцептуальні міри якості PESQ та PEAQ є набагато складнішими, порівняно із (1)-(4), мірами, оскільки базуються на комп'ютерному моделюванні слухової системи людини. Суттєвим недоліком вказаних та багатьох інших мір є високий рівень складності відповідного програмного забезпечення. Тому доцільним було звернути особливу увагу на міри (1) та (2), які хоча і є простими для обчислень, проте все ще є недостатньо вивченими.

Автором досліджено чутливість міри SSNR до похибки вирівнювання музичних сигналів за часом. На графіках рис. 3 наведено експериментально отримані залежності DMOS(Δf) для суб'єктивного оцінювання якості музичних сигналів, спотворених через обмеження смуги частот Δf , де граничні лінії відповідають 95%-му довірчому інтервалу. Аналіз обчислених за виразом (1) залежностей SSNR(Δf) (рис. 4) дозволив, по-перше, встановити, що вказана підвищена чутливість проявляється в порушенні монотонності залежності

$SSNR(\Delta f)$. По-друге, результати досліджень свідчать, що для нейтралізації зазначеного недоліку необхідно підвищувати частоту дискретизації сигналів, що порівнюються, не менше ніж в 4-5 разів, порівняно із частотою дискретизації запису. Отримані результати добре узгоджуються з результатами попередніх аналогічних досліджень властивостей міри $SSNR$ для мовних сигналів, де було показано, що при обчисленні $SSNR$ частоту дискретизації потрібно підвищувати щонайменше в 2 рази.

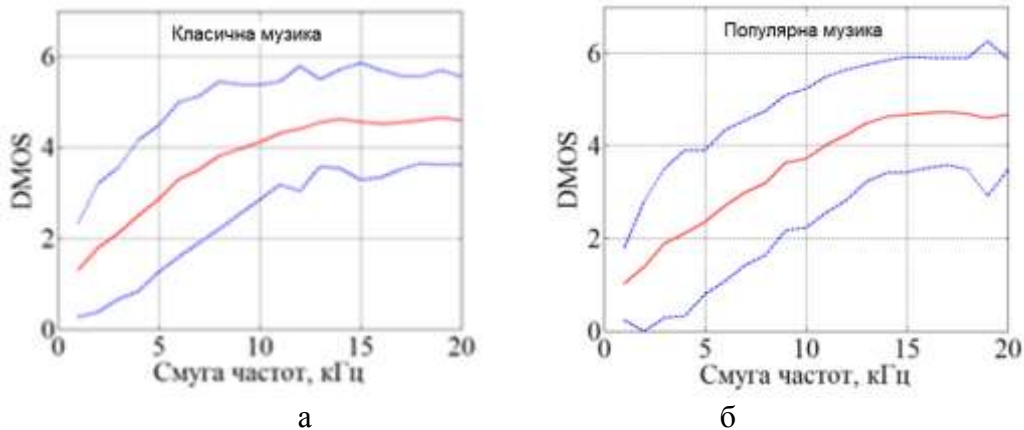


Рис. 3. Залежність $DMOS(\Delta f)$ для класичної (а) та популярної (б) музики

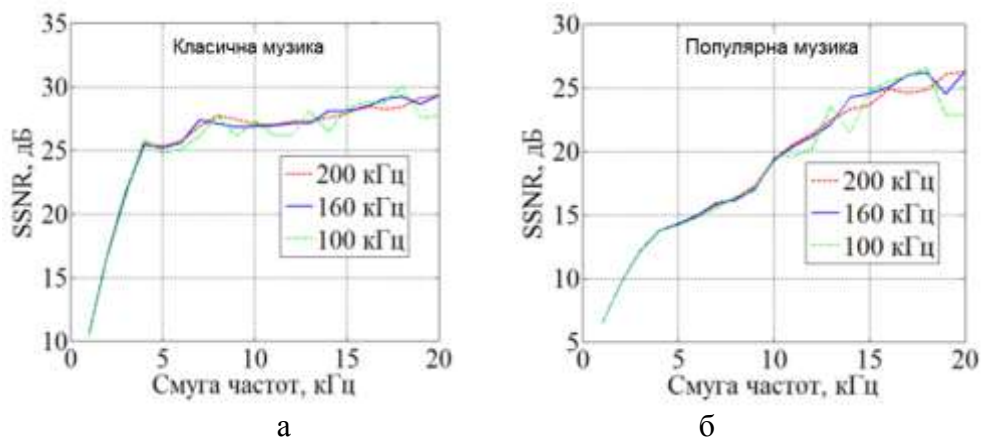


Рис. 4. Залежність $SSNR(\Delta f)$ для класичної (а) та популярної (б) музики

Аналогічні до $SSNR(\Delta f)$ залежності були отримані також для мір (2), (3) та $PEAQ$. Оцінки кореляції між зазначеними об'єктивними мірами та суб'єктивною мірою $DMOS$ наведено в табл. 1.

Таблиця 1. Коефіцієнти кореляції оцінок якості

Об'єктивний показник	$SSNR$	LSD	BSD	$PEAQ$
Коефіцієнт кореляції	0.94	-0.93	-0.92	0.94

Згідно літературних джерел, обчисленням таких коефіцієнтів кореляції дослідники часто завершують оцінювання достовірності об'єктивних мір якості сигналів, що є недостатньо коректним, оскільки таким чином не припускається можливість нелінійної залежності між результатами об'єктивної та суб'єктивної експертизи. Усунути вказаний недолік можна шляхом побудови карт відповідності результатів об'єктивної та суб'єктивної експертизи. Такі

карти є необхідними для практичних застосувань, оскільки дозволяють здійснити калібрування шкал об'єктивних мір якості.

На рис. 5 наведено отримані автором карти відповідності для мір SSNR та PEAQ, з яких видно, що в інтервалі DMOS=1-4,5 проста для обчислень міра SSNR може бути хорошою альтернативою складній мірі PEAQ.

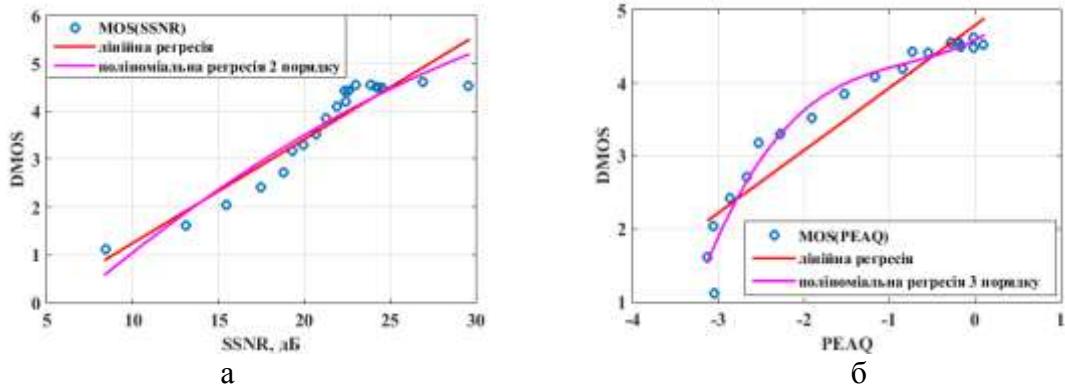


Рис. 5. Карти відповідності для SSNR (а) та PEAQ (б)

При дослідженні властивостей об'єктивної міри LSD в задачі оцінювання впливу обмеженості смуги частот Δf на якість музичних сигналів виявилося, що залежності $LSD(\Delta f)$ можуть також містити порушення монотонності (рис. 6).

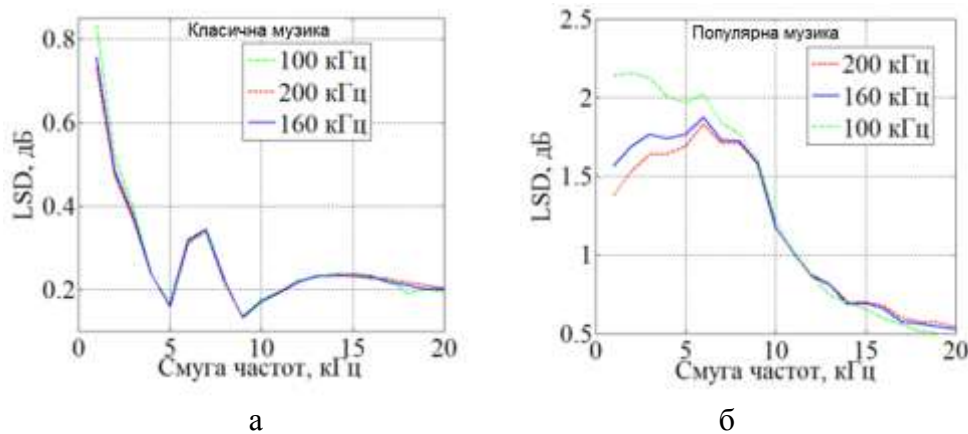


Рис. 6. Залежність $LSD(\Delta f)$ для класичної (а) та популярної (б) музики

Аналіз цього явища показав існування таких його причин: 1) «природня» висока чутливість міри LSD до сплесків в спектрі процесу; 2) нормування сигналів, що порівнюються, за дисперсією; 3) відхилення форми амплітудної частотної характеристики (АЧХ) НЧ фільтрів від прямокутної форми.

Для досліджень впливу фазових спотворень на якість мовних та музичних сигналів використано комп'ютерну модель системи із нелінійною фазочастотною характеристикою (ФЧХ), яка містить гребінку із семи смугових нерекурсивних октавних фільтрів із центральними частотами f_0 від 125 Гц до 8 кГц, лінії затримки та суматор (рис. 7). Як міру ступеня нелінійності ФЧХ

$\theta(f)$ використано груповий час затримки $\tau(f) = -\frac{1}{2\pi} \frac{d\theta(f)}{df}$. Тип ФЧХ, для

якого залежність $\tau(f)$ є спадаючою функцією, було позначено ФЧХ1, тоді як тип ФЧХ, де $\tau(f)$ є зростаючою функцією, було позначено ФЧХ2.

Результати суб'єктивного оцінювання якості мовних та музичних сигналів (рис. 8) свідчать, що слухова система людини є більш чутливою до другого виду фазових спотворень (ФЧХ2). При цьому об'єктивні оцінки (1)-(4), PESQ та PEAQ виявилися не чутливими до різниці між ФЧХ1 та ФЧХ2.

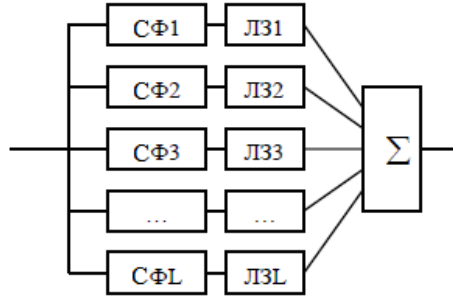


Рис. 7. Гребінка смугових фільтрів із лініями затримки та сумматором

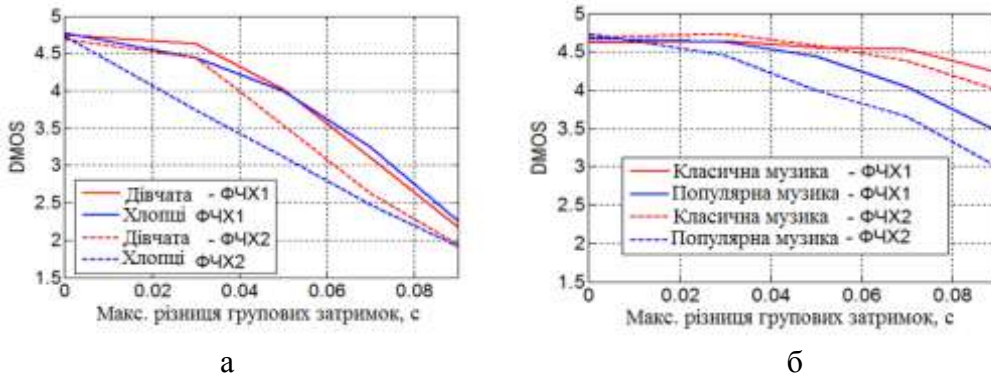


Рис. 8. Суб'єктивні оцінки якості мови (а) та музики (б)

Ці результати також свідчать, що фазові спотворення мовних сигналів суб'єктивно і об'єктивно сприймаються сильніше, ніж для музичних сигналів:

$$\Delta\tau_{\max, DMOS \geq 4} \approx \begin{cases} 0.4 \text{ с} & \text{для мови;} \\ 0.8 \text{ с} & \text{для музики.} \end{cases}$$

Вплив нелінійних спотворень на якість мовних та музичних сигналів досліджено для випадку жорсткого двостороннього кліпування сигналу $x(n)$, коли спотворений сигнал $y(n)$ формується згідно рівняння

$$y(n) = \begin{cases} x(n), & |x| < A, \\ A \cdot \text{sign}[x(n)], & |x| \geq A, \end{cases}$$

де $\text{sign}(\cdot)$ - знакова функція; A - рівень обмеження амплітуди сигналу.

Залежності суб'єктивних оцінок якості мовних та музичних сигналів від ступеня спотворення $k = 20 \lg(\max |x(n)|/A)$ наведені на рис. 9.

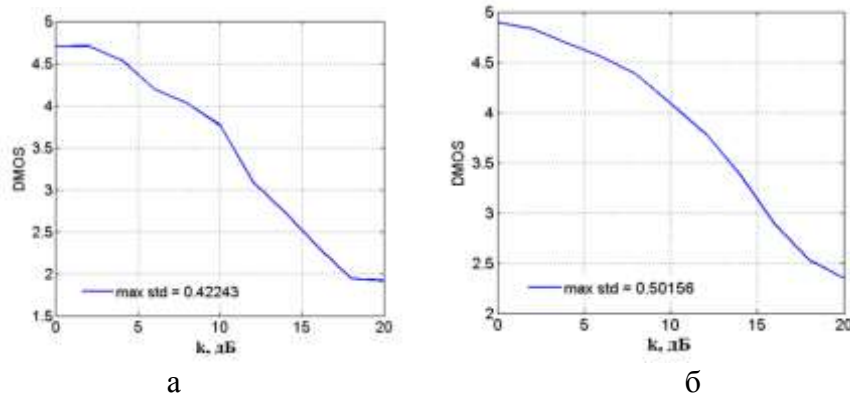


Рис. 9. Залежності $DMOS(k)$ для мови (а) та музики (б)

Як об'єктивну міру якості кліпованих сигналів запропоновано використовувати ексцес (пов'язаний із коефіцієнтом ексцесу k виразом $k = \beta_4 - 3$)

$$\beta_4 = \frac{\mu_4}{(\mu_2)^2},$$

де μ_k – центральний момент k -го порядку. Крім того, запропоновано дві нові міри ступеня кліпування, такі як величина, обернена до ексцесу, а також корінь квадратний із величини, оберненої до ексцесу:

$$\gamma_4 = 1/\beta_4, \quad \eta_4 = \sqrt{\gamma_4} = 1/\sqrt{\beta_4}.$$

Особливо зручною для практичного використання виявилася міра $\eta_4(k)$, яка є нормованою дисперсією сигналу. Її карти відповідності є практично лінійними функціями (рис. 10), що спрощує калібрування шкали значень η_4 .

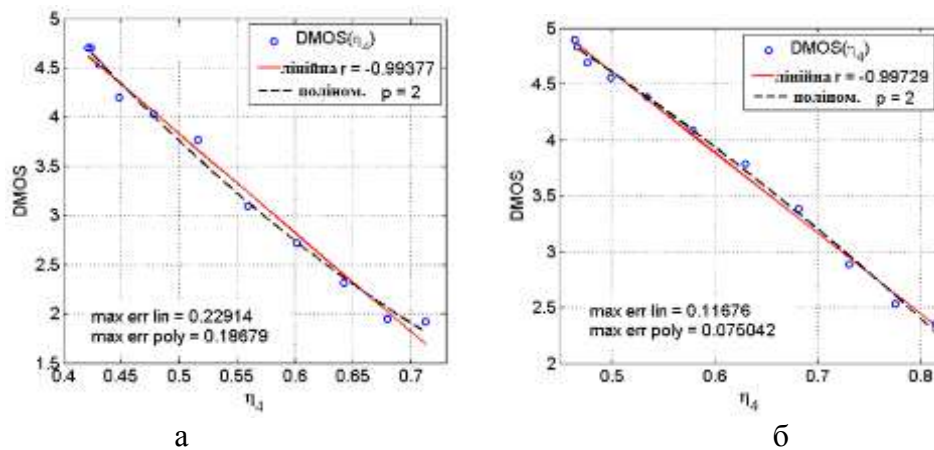


Рис. 10. Карти відповідності $DMOS(\eta_4)$ для мови (а) та музики (б)

У третьому розділі представлено результати, отримані за допомогою розробленої автоматизованої системи суб'єктивного оцінювання розбірливості мови в умовах дії шумової та ревербераційної завад, а також досліджено можливість об'єктивного оцінювання розбірливості мови шляхом оцінювання об'єктивних мір якості мовних сигналів. Узагальнену блок-схему керуючої програми автоматизованої системи суб'єктивного оцінювання розбірливості мови наведено на рис. 11. Моделювання сигналів, спотворених адитивним шумом та реверберацією, виконувалося за алгоритмом:

$$y(t) = x(t) \otimes h(t) + n(t)$$

де $x(t)$ – мовний сигнал, $h(t)$ – імпульсна характеристика приміщення, $n(t)$ – адитивний шум, \otimes – символ згортки. Цей алгоритм дозволяє легко врахувати особливості структури імпульсної характеристики приміщення:

$$h(t) = h_e(t) + h_l(t), \quad h_e(t) = \begin{cases} h(t), & t \in 0..50 \text{ ms}; \\ 0, & t \notin 0..50 \text{ ms}, \end{cases} \quad h_l(t) = \begin{cases} h(t), & t > 50 \text{ ms}; \\ 0, & t \leq 50 \text{ ms}, \end{cases}$$

де $h_e(t)$ описує ранні відбиття, а $h_l(t)$ описує пізню реверберацію.

За допомогою розробленої автоматизованої системи виконано порівняння оцінок розбірливості мови, отриманих для умов прослуховування через навушники (рис. 12) та акустичні монітори.

Зроблено важливий для практичного використання висновок щодо

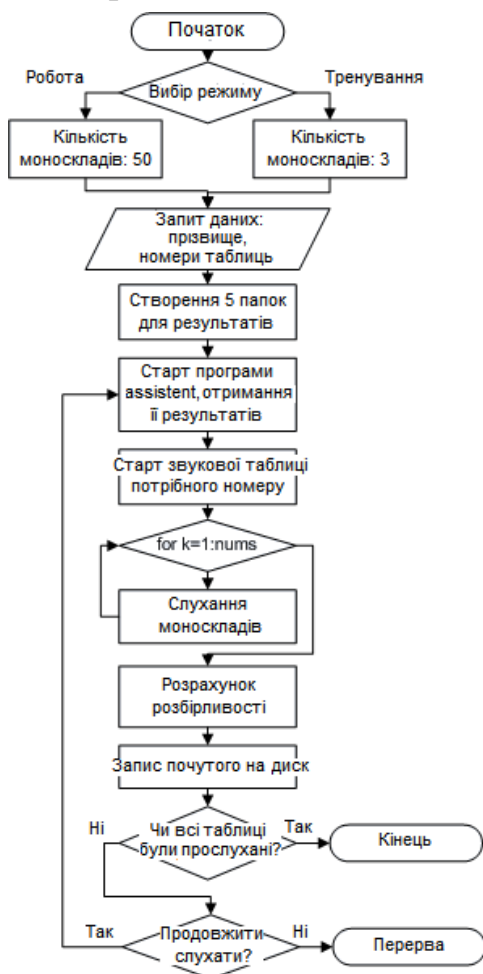


Рис. 11. Блок-схема керуючої програми

можливості оцінювання розбірливості мови навіть із акустичними моніторами в звичайному приміщенні, але за умов, що дистанція до моніторів не перевищує так звану критичну дистанцію (для побутових та лекційних приміщень критична дистанція є близькою до 0,8-1 м).

Експериментальне оцінювання розбірливості мови при бінауральному прослуховуванні (рис. 13) дозволило отримати нові результати. Кількісно оцінено ступінь переважаючого впливу шуму, порівняно із реверберацією, на розбірливість мови. При цьому отримано прямі оцінки розбірливості мови, що є більш зручним для практичних застосувань, ніж відношення сигнал-шум, яке використовувалося попередниками.

Крім того, представлені на рис. 13 результати свідчать, що негативний вплив реверберації є найбільшим всередині приміщення (порівняно із позиціями «поблизу диктора» та «поблизу стінки»).

Більш того, ці результати свідчать, що ранні відбиття не є такими корисними, як вважалося

раніше, коли було запропоновано коефіцієнт C50 як об'єктивна міра розбірливості мови. Насправді (рис. 13,б), ранні відбиття на 3-5% знижують розбірливість мови для слухачів, що віддалені від диктора та стін, що можна пояснити зниженням якості сигналу через нерівномірність АЧХ приміщення.

В даному розділі також досліджено принципову можливість об'єктивного оцінювання розбірливості мови непрямым шляхом – через об'єктивне оцінювання якості мовних сигналів.

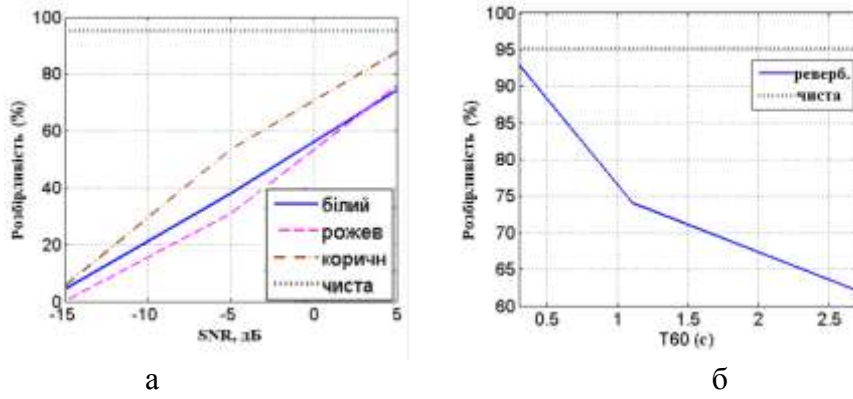


Рис. 12. Розбірливість мови: слухання через навушники для шуму (а) та реверберації (б)

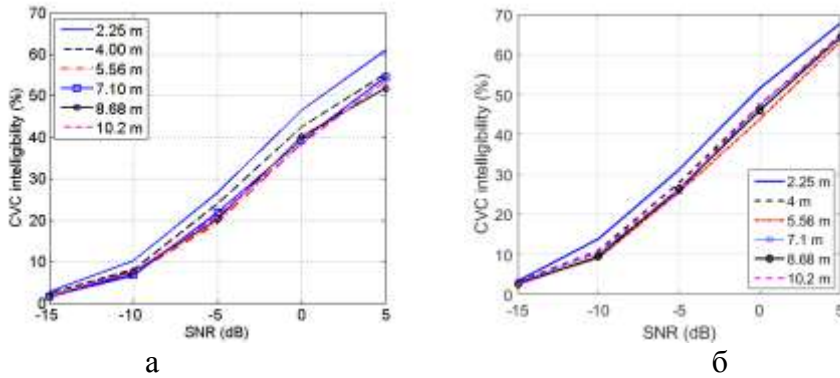


Рис. 13. Розбірливість мови: бінауральне слухання для сумісної дії шуму та реверберації (а) та для сумісної дії шуму та ранніх відбиттів (б)

Обґрунтуванням актуальності такої задачі є той факт, що оцінювання розбірливості мови об'єктивними методами є набагато складнішим, порівняно з оцінюванням якості мови. Дійсно, при використанні формантного методу словесну розбірливість W обчислюють через формантну розбірливість A :

$$W = \begin{cases} 1.54 \cdot A^{0.25} [1 - \exp(-11 \cdot A)], & A < 0.15; \\ 1 - \exp\left(\frac{11 \cdot A}{1 + 0.7 \cdot A}\right), & A \geq 0.15. \end{cases} \quad A = \sum_{k=1}^K p_k \cdot P(E'_k),$$

де $P_k(E'_k)$ – коефіцієнт сприйняття мови; $E_k \approx SNR_k = 10 \lg \frac{D_{sk}}{D_{nk}}$ – відношення сигнал-шум в k -й смузі частот; p_k – ймовірність перебування формант в k -й смузі частот з центральними частотами f_{0k} та граничними частотами f_{1k} та f_{2k} :

$$p_k = F(f_{2k}) - F(f_{1k}), \quad F(f) = \begin{cases} 2.57 \cdot 10^{-8} \cdot f^{2.4}, & 100 < f \leq 400 \text{ Гц}; \\ 1 - 1.074 \cdot \exp(-10^{-4} \cdot f^{1.18}), & 400 < f \leq 10000 \text{ Гц}. \end{cases}$$

Не менш складним для практичного використання є модуляційний метод, де тестовий сигнал $x(t)$ має вигляд модульованого за гармонічним законом нестационарного випадкового процесу, а мірою розбірливості є індекс STI.

На жаль, досі залишається відкритим питання, за яких умов можливе використання простих мір якості замість складних мір розбірливості. Окремі результати проведених в роботі експериментальних досліджень, спрямованих

на усунення зазначеного недоліку, наведені на рис. 14 та свідчать що для середніх значень відношення сигнал-шум ($\text{SNR} > 0$ дБ) дійсно можна значно спростити процедуру оцінювання розбірливості мови, замінивши її оцінюванням якості мовних сигналів.

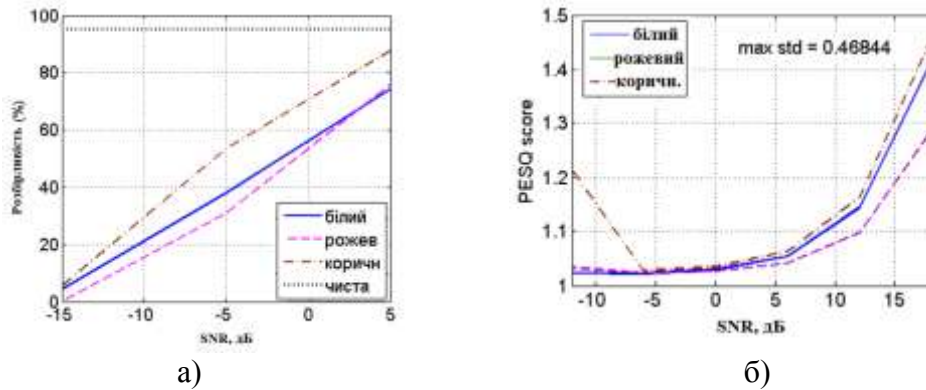


Рис. 14. Усереднені оцінки розбірливості мови (а) і якості мови (показник PESQ) (б).

Четвертий розділ присвячено порівнянню алгоритмів DS та LCMV просторово-часової обробки за критерієм якості сигналу на виході мікрофонного масиву (рис. 15). Найпростішим алгоритмом просторово-часової обробки сигналів з виходів окремих мікрофонів є алгоритм сум та затримок (DS-алгоритм):

$$y(n, \theta_t) = \sum_{m=1}^M x_m(n - \Delta_m(\theta_t)),$$

де $x_m(t)$ - сигнал на виході m -го мікрофона; Δ_m - затримка, розрахована відповідно напрямку надходження сигналу θ_t . Відомим недоліком такого алгоритму є залежність діаграми спрямованості від частоти. Усунути цей недолік та забезпечити підвищену гостроту діаграми спрямованості (ДС) можна забезпечити шляхом використання більш складного алгоритму

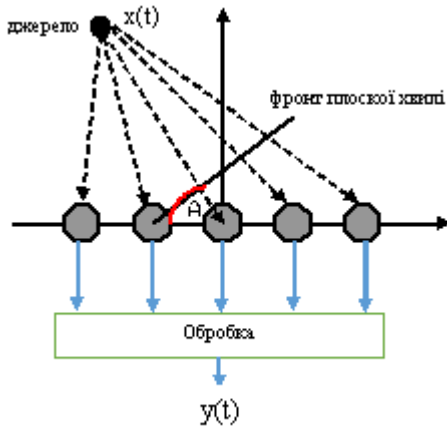


Рис. 15. Мікрофонний масив

обробки:

$$y(t, \theta_t) = \sum_{m=1}^M w_m(\theta_t, \theta_t) x_m(t - \Delta_m(\theta_t)),$$

де Δ_m затримка, розрахована відповідно напрямку надходження сигналу θ_t , w_m – вагові коефіцієнти, які залежать від частоти та напрямку θ_t . Зокрема, для алгоритму LCMV рівняння для функції вагових коефіцієнтів мають вигляд:

$$w_m(\theta_t, \theta_t) = (R_{xx} + \varepsilon I)^{-1} C [C^H (R_{xx} + \varepsilon I)^{-1} C]^{-1} C^H w$$

де R_{xx} – матриця просторової кореляції, C – матриця обмежень, I – одинична матриця, ε – параметр регуляризації. Наявність параметру регуляризації свідчить про можливість негативних наслідків такої обробки у вигляді спотворень форми результуючого сигналу $y(t)$.

Для перевірки існування та серйозності зазначеного недоліку, виконано експериментальні дослідження, де розглянуто чотири ситуації розташування джерела шуму:

- перша, друга і третя ситуації – одиночне джерело розмовного шуму (розмова групи людей), розташоване відповідно під кутами 15, 45 та 90 градусів до джерела музичного сигналу;
- четверта ситуація – два джерела розмовного шуму, розташованих під кутами до джерела музичного сигналу 45 та 120 градусів.

На рис. 16 наведено результати суб'єктивного оцінювання, за шкалою DMOS, якості музичних сигналів для алгоритмів LCMV та DS. По осі абсцис відкладено номери типу обробки: 1 – обробка відсутня (вхідний сигнал); 2 – сигнал з виходу одиночного мікрофона; 3 – сигнал з виходу DS-алгоритму; 4,5,6 – сигнал з виходу різновидів LCMV- алгоритму. Наведені графіки свідчать, що алгоритм LCMV помітно поступається алгоритму DS за якістю сформованих сигналів. Причиною виявленого ефекту є підвищена чутливість алгоритму LCMV до вибору розміру сегментів при посегментній обробці сигналів, а також до особливостей амплітудно-частотної характеристики мікрофонів.

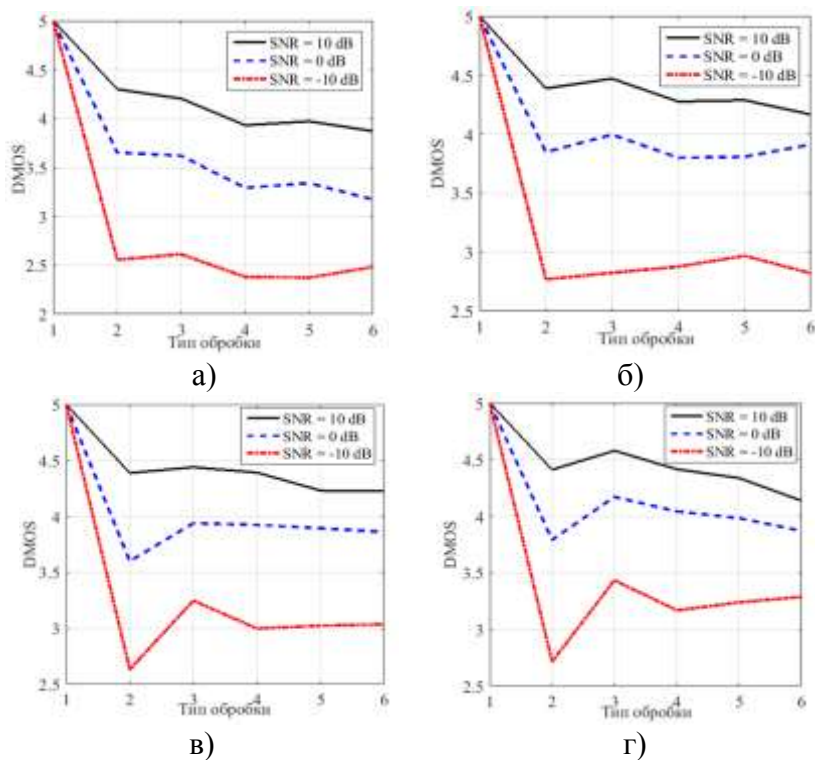


Рис. 16. Суб'єктивні оцінки якості сигналу для ситуації 1 (а), ситуації 2 (б), ситуації 3 (в) та ситуації 4 (г)

З наведеними на рис. 16 результатами суб'єктивного оцінювання добре узгоджуються результати об'єктивного оцінювання якості музичних сигналів, зокрема, отримані із використанням міри PEAQ. Таким чином, за результатами проведених досліджень, складний для обчислень алгоритм LCMV недоцільно використовувати замість простого алгоритму DS, оскільки зазначена складність обчислень не компенсується підвищеною якістю вихідного сигналу.

ВИСНОВКИ

Основні результати дисертаційної роботи полягають у наступному.

1. Вперше для реальних акустичних сигналів отримано кількісні суб'єктивні та об'єктивні оцінки ступеню впливу нелінійності фазової характеристики тракту на якість мовних та музичних сигналів, що дозволяє сформулювати адекватні вимоги до ступеня спотворення акустичних сигналів, а саме:

- оскільки фазові спотворення музичних сигналів значно менш помітні на слух, ніж такі для мовних сигналів, для якісного відтворення музичних сигналів різниця часу групової затримки на низьких та високих частотах не має перевищувати 80 мс, а мовних сигналів – 40 мс;

- слухова система людини є чутливою до знаку різниці часу групової затримки для низьких та високих частот, що має враховуватися при налагодженні систем відтворення мовних та музичних сигналів.

2. Вперше запропоновано використання кумулянтного коефіцієнту статистичного розподілу миттєвих значень акустичного сигналу у вигляді коефіцієнта ексцесу та його функціональних перетворень як об'єктивних мір ступеня кліпування та якості мовних та музичних сигналів, спотворених кліпуванням. Дана пропозиція дозволяє суттєво (в 200-300 разів) знизити вимоги до розмірності вектору параметрів, що оцінюється при виявленні кліпування і таким чином значно скоротити час відповідних обчислень без зниження достовірності результатів.

3. Вперше вказано на необхідність врахування підвищеної чутливості простої в обчислювальному плані об'єктивної міри якості сигналів у вигляді лог-спектральних спотворень до спектрального складу акустичних сигналів, що аналізуються. Отримані рекомендації з урахування такої підвищеної чутливості дозволяють підвищити достовірність висновків при оцінюванні якості спотворених мовних та музичних сигналів.

4. Знайшло подальший розвиток положення про принципову можливість підвищення достовірності об'єктивного оцінювання якості спотворених акустичних сигналів з використанням такої міри як сегментне відношення сигнал-шум та показано, що такий недолік як чутливість сегментного відношення сигнал-шум до похибки вирівнювання сигналів може бути нейтралізовано збільшенням частоти дискретизації шляхом інтерполяції щонайменше в 5 разів для музичних сигналів, тоді як для мовних сигналів частоту дискретизації достатньо збільшувати в 2 рази.

5. Удосконалено технологію автоматизації суб'єктивного оцінювання розбірливості мови, спотвореної шумом та реверберацією, артикуляційним методом, для чого запропоновано при фіксації почутих односкладових слів записувати повне слово замість дописування відсутніх літер, а також розроблено та протестовано відповідне програмне забезпечення. Запропонована технологія є значно простішою у використанні, порівняно із зарубіжними

аналогами, що дозволяє значно пришвидшити та знизити вартість проведення артикуляційних досліджень без зниження їх достовірності.

6. Знайшло подальший розвиток положення про межі застосування певного переліку об'єктивних оцінок якості мови замість суб'єктивних оцінок розбірливості зашумленої мови, що дозволяє при значеннях відношення сигнал-шум, більших за 0 дБ, оцінювати розбірливість мови через оцінювання якості мови й таким чином значно пришвидшити процедуру об'єктивного оцінювання розбірливості мови без зниження достовірності оцінок.

7. Знайшло подальший розвиток положення про значну роль ранніх відбиттів звуку в приміщеннях, для чого виконано порівняння результатів артикуляційних досліджень при монауральному та бінауральному режимах прослуховування мовних сигналів, спотворених шумом та реверберацією. Отримані результати свідчать, що ранні відбиття хоча і є корисними з енергетичної точки зору, оскільки призводять до збільшення відношення сигнал-шум на 6-9 дБ, проте вони є шкідливими з інформаційної точки зору, оскільки розбірливість мови знижується на 3-5% на відстанях 5-9 м до джерела звуку, що спричинене нерівномірністю частотної характеристики приміщення.

8. Знайшло подальший розвиток зіставлення, шляхом суб'єктивного та об'єктивного оцінювання якості звуку для кількох алгоритмів обробки спотворених шумом музичних сигналів в багатомікрофонних масивах та показано, що при використанні простішого, з обчислювальної точки зору, алгоритму "сум та затримок" можна отримати більш якісні сигнали, ніж при використанні складних алгоритмів, що забезпечують незалежні від частоти діаграми підвищеної спрямованості.

9. Проведено попередній аналіз причин погіршеної якості сигналів для алгоритмів, що забезпечують незалежні від частоти діаграми підвищеної спрямованості, та вказано, що таке погіршення спричинене значною похибкою обчислень при реалізації цих алгоритмів, причому ця похибка зростає при зменшенні довжини сегментів сигналу та при зменшенні смуги пропускання мікрофонів відносно частоти дискретизації. Отримані результати мають сприяти подальшим дослідженням та застосуванню багатомікрофонних систем просторово-часової обробки звукових сигналів, спотворених шумом та реверберацією.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. И.В. Котвицкий, А.Н. Продеус, «Объективное и субъективное оценивание качества речевых и музыкальных сигналов, подвергнутых фазовым искажениям». - ISSN 1811-4512. *Электроника и связь*. 2016, № 2. С. 25-31. *Особистий внесок здобувача полягає в комп'ютерному моделюванні фазових спотворень (Index Copernicus)*.
2. A.M. Prodeus, I.V. Kotvytskyi, A.A. Ditiashov, «Assessment of clipped speech quality», *Electronics and Control Systems*, 2018, N 4(58), P. 11-18. DOI:10.18372/1990-5548.58.13504. *Особистий внесок здобувача полягає в*

- проведенні розрахунків, організації експерименту та аналізі результатів експерименту (*Index Copernicus*).
3. А. Продеус, А. Вітик, О. Дворник, І. Котвицький, О. Чайка, М. Ярошенко, «Суб'єктивне оцінювання розбірливості мови на тлі шуму та реверберації». Мікросистеми, електроніка та акустика. Том 23, № 2, 2018. ISSN 2523-4447, DOI: 10.20535/2523-4455.2018.23.2.128820. С. 66-73. *Особистий внесок здобувача полягає в дослідженні комп'ютерної моделі системи артикуляційних досліджень (Index Copernicus).*
 4. М.П. Юраков, І.В. Котвицький, А.М. Продеус, «Оцінювання якості музичних сигналів, обмежених за смугою частот». ISSN 2523-4447. Мікросистеми, Електроніка та Акустика, 2018, vol. 23, № 4,. DOI: 10.20535/2523-4455.2018.23.4.133785. С. 58-64. *Особистий внесок здобувача полягає в проведенні комп'ютерного моделювання, проведенні експерименту та аналізі результатів експерименту (Index Copernicus).*
 5. A. Prodeus, K. Bukhta, P. Morozko, O. Serhienko, I. Kotvytskyi, O. Dvornyk, «Automated Subjective Assessment of Speech Intelligibility in Various Listening Modes», ISSN 2523-4447. Microsystems, Electronics and Acoustics, 2018, vol. 23, no. 3, P.49-57. *Особистий внесок здобувача полягає в дослідженні комп'ютерної моделі системи артикуляційних досліджень та проведенні експерименту (Index Copernicus).*
 6. I. Kotvytskyi, «Objective and Subjective Music Quality Assessment for Signals Limited by Frequency Band». IEEE International Young Scientists Forum on Applied Physics and Engineering (YSF). м. Львів, 16-18 жовтня 2017 року.
 7. А.Н. Продеус, В.С. Дидковский, И.В. Котвицкий, «Об оценивании качества речи с помощью показателя «логарифмические спектральные искажения»». Тези доповідей 5-ї міжнародної науково-технічної конференції «Інформаційні проблеми теорії акустичних, радіоелектронних і телекомунікаційних систем» IPST-2016. – Харків: НТУ «ХПІ», 2016. С. 14. *Особистий внесок здобувача полягає у розрахунку та аналізі об'єктивних мір якості сигналів.*
 8. A. Prodeus, V. Didkovskyi, M. Didkovska, I. Kotvytskyi, «Automated Subjective Evaluation of Speech Intelligibility in Noise and Reverberation». *Proceedings of IEEE 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, February 20-24, 2018, Lviv-Slavske, Ukraine, P. 555-558. ISBN: 978-1-5386-2557-6. DOI: 10.1109/TCSET.2018.8336263 *Особистий внесок здобувача полягає в дослідженні комп'ютерної моделі системи артикуляційних досліджень та проведенню експерименту.*
 9. A. Prodeus, I. Kotvytskyi, «On Reliability of Log-Spectral Distortion Measure in Speech Quality Estimation», *Proceedings of IEEE 5th International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD)*, P. 17-19. October 2017, Kyiv, Ukraine. *Внесок здобувача полягає у розрахунку та аналізі об'єктивних мір якості сигналів.*

10. A. Prodeus, V. Didkovskiy, M. Didkovska, I. Kotvytskyi, «On Peculiarities of Evaluating the Quality of Speech and Music Signals Subjected to Phase Distortion», *Proceedings of IEEE 37th International Conference on Electronics and Nanotechnology (ELNANO)*, April 18-20, 2017, Kyiv, Ukraine. – P. 455-460. *Особистий внесок здобувача полягає в розробці та дослідженні комп'ютерних моделей.*
11. A. Prodeus, I. Kotvytskyi, A. Ditiashov, «Clipped Speech Signals Quality Estimation», *Proceedings of 5th International Conference "Methods and Systems of Navigation and Motion Control" (MSNMC-2018)*, October 16-18, 2018, Kyiv, Ukraine, P. 151-155. ISBN: 978-1-5386-5869-7. DOI: 10.1109/MSNMC.2018.8576308. *Особистий внесок здобувача полягає в проведенні комп'ютерного моделювання та проведенні експерименту.*
12. A. Prodeus, K. Bukhta, P. Morozko, O. Serhiienko, I. Kotvytskyi, I. Shherbenko, «Automated System for Subjective Evaluation of the Ukrainian Speech Intelligibility», *Proceedings of IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, April 24-26, 2018 Kyiv, Ukraine, P. 533-538. ISBN: 978-1-5386-6382-0. *Особистий внесок здобувача полягає в проведенні комп'ютерного моделювання.*
13. A. Prodeus, I. Kotvytskyi, M. Didkovska, «Kurtosis and Its Transformations as Objective Measures of Clipping Value and Speech Quality», *Proc. of IEEE 5th Int. Conf. Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD)*, P. 22-24 October 2019, Kyiv, Ukraine. *Особистий внесок здобувача полягає в проведенні комп'ютерного моделювання.*
14. A. Prodeus, M. Didkovska, I. Kotvytskyi, «Using Kurtosis for Objective Assessment of the Musical Signals Clipping Degree», *Proc. 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, October 2019, Kyiv, Ukraine. *Особистий внесок здобувача полягає в проведенні комп'ютерного моделювання.*
15. A. Prodeus, I. Kotvytskyi, A. Garasiuk, A. Semenchuk, A. Telenga, S. Klymenko, V. Lozinskyi, O. Andriichenko, O. Denysenko, «Subjective Assessment of the Influence of Room Characteristics on Intelligibility of Noised Speech». *Proceedings of IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO)*, P. 16-18 April 2019, Kyiv, Ukraine. DOI: 10.1109/ELNANO.2019.8783250. *Особистий внесок здобувача полягає в дослідженні комп'ютерної моделі системи артикуляційних досліджень та проведенні експерименту.*
16. Комп'ютерна програма «Автоматизована система суб'єктивного оцінювання розбірливості мови»: свідоцтво про реєстрацію авторського права на твір № 85027/І. В. Котвицький та ін. Україна: Міністерство економічного розвитку і торгівлі України; опубл. 30.01.19. *Особистий внесок здобувача полягає у тестуванні автоматизованої системи оцінювання розбірливості мови.*

АНОТАЦІЯ

Котвицький І. В. Оцінювання якості спотворених мовних та музичних сигналів. - На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.09.08 - прикладна акустика та звукотехніка. – Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», МОН України, Київ, 2020.

Дисертаційна робота присвячена об'єктивізації та автоматизації оцінювання якості мовних та музичних сигналів, спотворених в комунікаційних каналах. Побудовано карти відповідності між суб'єктивною мірою якості DMOS та об'єктивними мірами якості акустичних сигналів. Вперше запропоновано об'єктивні міри ступеня кліпування мовних та музичних сигналів. Вдосконалено технологію автоматизації суб'єктивного оцінювання розбірливості мови. Знайшло подальший розвиток зіставлення, шляхом об'єктивного та суб'єктивного оцінювання якості звуку, алгоритмів обробки музичних сигналів в багатомікрофонних масивах.

Ключові слова: карта відповідності, кліпований сигнал, масив мікрофонів, міра якості, реверберація, розбірливість мови, суб'єктивна оцінка, фазові спотворення, шум, якість сигналу.

АННОТАЦИЯ

Котвицкий И. В. Оценивание качества искаженных речевых и музыкальных сигналов. - На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.09.08 - прикладная акустика и звукотехника. - Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского», МОН Украины, Киев, 2020.

Диссертация посвящена объективизации и автоматизации оценивания качества речевых и музыкальных сигналов, искаженных в коммуникационных каналах. Построены карты соответствия субъективной меры качества DMOS и объективных мер качества акустических сигналов. Впервые предложены объективные меры степени клипирования акустических сигналов. Усовершенствована технология автоматизации субъективного оценивания разборчивости речи. Дальнейшее развитие получило сопоставление, путем объективного и субъективного оценивания качества звука, алгоритмов обработки музыкальных сигналов в многомикрофонных массивах.

Ключевые слова: карта соответствия, качество сигнала, клипированный сигнал, массив микрофонов, мера качества, разборчивость речи, реверберация, субъективная оценка, фазовые искажения, шум.

SUMMARY

Kotvytskyi I. V. Assessing the quality of distorted speech and music signals.

- Manuscript.

Thesis for a Candidate degree in Technical Science, Specialty 05.09.08 Applied Acoustics and Audio Engineering. - National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ministry of Education and Science of Ukraine, Kyiv, 2020.

The thesis is devoted to the objectification and automation of evaluation of the quality of speech and music signals distorted in communication channels. Particular attention is paid to the analysis of the objective measures of the quality of speech and music signals. They are segmental signal-to-noise ratio SSNR, log-spectral distortion LSD, bar-spectral distortion BSD, perceptual assessment of the speech quality PESQ and perceptual evaluation of the music quality PEAQ. These measures are promising for use in real-time signal quality assessment and correction systems. The matching maps between subjective DMOS quality measure and objective SSNR, LSD, BSD and PEAQ measures for situations of linear and non-linear signal distortion has been proposed. At the same time, such types of linear distortions as limiting the frequency band and nonlinearity of the phase frequency response of the communication channel were considered. The necessity of taking into account the increased sensitivity of the simple measure in the computational plan of an objective measure of the quality of signals in the form of log-spectral distortions to the spectral composition of the acoustic signals being analyzed is pointed out. The received recommendations for such high sensitivity make it possible to increase the reliability of conclusions when evaluating the quality of distorted speech and music signals. It further developed the provision on the fundamental possibility of increasing the reliability of objective estimation of the quality of distorted acoustic signals using such measures as the segmental signal-to-noise ratio, and it is shown that such a disadvantage as the sensitivity of the segmental signal-to-noise ratio to the error of signal alignment can be counteracted by increasing the sampling rate by interpolating at least 5 times for music signals, while for speech signals it is sufficient to increase the sampling rate by 2 times. To investigate the effects of phase distortions on the quality of speech and music signals, a computer model of a system with nonlinear frequency-phase response was used which contains a seven-band non-recursive octave filter, delay lines, and summation block. The group delay time was used as a measure of the nonlinearity of the frequency-phase response. Two types of group delay time has been considered: the first type is described by the decreasing dependence of the group delay time on the frequency, and the second, on the contrary, is described by the increasing dependence of the group delay time on the frequency. It was shown that the human auditory system does not perceive phase distortion, provided that the difference in the group delay times at high and low frequencies does not exceed 40 ms for speech signals and 80 ms for music signals. In addition, it was found that the human auditory system is sensitive to the sign of the difference in group delay times

at high and low frequencies. The dependences of subjective quality indicators on clipping of speech and musical signals on the value of their clipping provide the necessary basis for further calibration of objective measures of the clipping degree. At the same time, objective measures of the clipping value of speech and musical signals has been proposed. These measures are kurtosis, inverse to kurtosis, and the square root of inverse to kurtosis. The results make it possible to calibrate the corresponding software and hardware systems for objective evaluation of the quality of distorted speech and music signals. The limits of the use of objective speech quality assessments instead of subjective noisy speech intelligibility have been clarified, which can greatly simplify the procedure for objectively assessing speech intelligibility. The technology of automation of subjective assessment of speech intelligibility by the articulation method has been improved and its efficiency was tested by means of the developed prototype of the automated system of articulation testing. The results of comparison, by objective and subjective assessment of sound quality, of signals processing algorithms in the multi-microphone arrays are refined. Comparison of two algorithms of space-time processing based on the criterion of the quality of the music signal distorted by additive interference allows us to recommend a simpler, from the point of view of calculations, algorithm of the sum and delays instead of the more complicated competitive algorithm. The analysis of the causes of the poor signal quality for algorithms providing frequency-independent diagrams of the increased directivity is analyzed, and it is stated that such deterioration is caused by a significant error of calculations in the implementation of these algorithms, and this error increases with decreasing the length of the signal segments and decreasing the microphone's frequency band. Thus, the results obtained indicate that microphone arrays, although they can be used to improve the quality of acoustic signals distorted by noise and reverberation, however, should be taken into account the increased sensitivity of the respective algorithms for signal processing to calculation errors. The results obtained by the author can be used in the acoustic examination of rooms and communication lines, as well as in the educational process of colleges.

Keywords: clipped signal, matching map, microphone array, noise, phase distortion, quality measure, reverberation, signal quality, speech intelligibility, subjective evaluation.